AD-A105 415 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/1
A SURVEY OF SOME PROBLEMS AND RECENT RESULTS FOR PARAMETER ESTI--ETC(U)
JUL 81 H T BANKS DAAG29-79-C-0161
UNCLASSIFIED LCDS-81-19 AFOSR-TR-81-0670 NL

1 OF 1
AD A
105415

END
DATE
FILMED
10-81
DTIC

Lefschetz Center for Dynamical Systems

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 81 -0670	2. GOVT ACCESSION NO.: AD-A105415
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): "A Survey of Some Problems and Recent Results for Parameter Estimation and Optimal Control in Delay and Distributed Parameter Systems"	5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL
	6. PERFORMING ORG. REPORT NUMBER: LCDS #81-19
7. AUTHOR(s): H. T. Banks	8. CONTRACT OR GRANT NUMBER(s): AFOSR 76-3092
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Lefschetz Center for Dynamical Systems, Div. of Appl.Math., Brown Univ., Prov. RI 02912	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PE61102F, 2304/A4
11. CONTROLLING OFFICE NAME AND ADDRESS: AFOSR, Bolling AFB, Washington, DC 20332	12. REPORT DATE: July 1981
	13. NUMBER OF PAGES: 28
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The investigator surveyed a number of applications and problems motivating his current efforts on numerical techniques for parameter estimation in and optimal control of delay and partial differential equations. He then outlined two different approaches for establishing theoretical convergence results for estimation algorithms. An application of modal techniques to the investigation of transport in brain tissue is briefly explained. A sketch of a convergence theory for spline techniques for function space parameter estimation problems is given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A SURVEY OF SOME PROBLEMS AND RECENT RESULTS FOR PARAMETER ESTIMATION AND OPTIMAL CONTROL IN DELAY AND DISTRIBUTED PARAMETER SYSTEMS

by

H. T. Banks

July, 1981 LCDS Report #81-19

A SURVEY OF SOME PROBLEMS AND RECENT RESULTS FOR PARAMETER ESTIMATION AND OPTIMAL CONTROL IN DELAY AND DISTRIBUTED PARAMETER SYSTEMS*+

Jul[redacted]81

H. T. Banks++

Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, RI 02912

*Invited lecture, Conference on Volterra and Functional Differential Equations, V.P.I.S.U., Blacksburg, VA, June 10-13, 1981.

+Work reported herein was supported in part by the Air Force Office of Scientific Research under contract AFOSR 76-3092[redacted], in part by the National Science Foundation under grant NSF-MCS 7905774-02, and in part by the U.S. Army Research Office under contract ARO-DAAG29-79-C-0161.

++Parts of the research discussed here were carried out while the author was a visitor at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA, which is operated under NASA contracts No. NAS1-15810 and No. NAS1-16394. Additional research and the manuscript was completed while the author was a visitor in the Mathematics Department, University of Utah.

A SURVEY OF SOME PROBLEMS AND RECENT RESULTS FOR PARAMETER ESTIMATION AND OPTIMAL CONTROL IN DELAY AND DISTRIBUTED PARAMETER SYSTEMS

by

H. T. Banks

ABSTRACT

We survey a number of applications and problems motivating our current efforts on numerical techniques for parameter estimation in and optimal control of delay and partial differential equations. We then outline two different approaches for establishing theoretical convergence results for estimation algorithms. An application of modal techniques to the investigation of transport in brain tissue is briefly explained. A sketch of a convergence theory for spline techniques for function space parameter estimation problems is given.

§1. Introduction.

In this lecture we shall first present a brief account of several areas of applications which have motivated our recent efforts, both theoretical and numerical, on approximation methods for estimation and control of infinite dimensional systems. We then shall sketch the general theoretical ideas we have employed to establish convergence results for related iterative schemes. Finally we return to two of the applications and illustrate the use of these ideas by explaining in more detail our investigations for these problems. As we shall make clear, our efforts on many of the problems mentioned below involve joint endeavors with colleagues and students. In addition to a well-deserved thank you to Richard Ambrasino, James Crowley, Patti Daniel, Mary Garrett, Karl Kunisch, and Gary Rosen, we would also like to publicly acknowledge E. Armstrong (NASA Langley Research Center), R. Ewing and G. Moeckel (Mobil Research and Development Corp.), P. Kareiva (Brown University), J. P. Kernevez (Universitè de Technologie de Compiègne), W. T. Kyner (University of New Mexico), and G. A. Rosenberg (V. A. Medical Center, U. N. M. School of Medicine) for numerous stimulating discussions and suggestions which have substantially affected the investigations of our group at Brown University.

Our discussions here focus on a general class of systems including nonlinear delay systems

$$(1)\qquad \begin{aligned} \dot{x}(t) &= f(\alpha,t,x(t),x_t,x(t-\tau_1),\ldots,x(t-\tau_\nu)) + g(t) \ , \\ x(\theta) &= \phi(\theta) \ , \ -\tau_\nu \leq \theta \leq 0 \ , \\ q &= (\alpha,\tau_1,\ldots,\tau_\nu) \ , \end{aligned}$$

nonlinear distributed parameter systems

$$\frac{\partial^2 u}{\partial t^2} = q_1 \frac{\partial^2 u}{\partial x^2} + q_2 \frac{\partial u}{\partial t} + q_3 u + f(q_6,t,x,u) ,$$

(2)
$$u(0,x) = q_4\phi(x) , \frac{\partial u}{\partial t}(0,x) = q_5\psi(x)$$

$$u(t,0) = g_1(t) , u(t,1) = g_2(t)$$

of hyperbolic type, and parabolic systems of the form

$$\frac{\partial u}{\partial t} = \frac{q_1}{k(x)} \frac{\partial}{\partial x} (p(x)\frac{\partial u}{\partial x}) + q_2 u + f(q_4,t,x,u) ,$$

(3)
$$u(0,x) = q_3\phi(x) ,$$

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \end{pmatrix} \left(u(t,0), \ \frac{\partial u}{\partial x}(t,0), \ u(t,1), \ \frac{\partial u}{\partial x}(t,1) \right)^T = \begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix} .$$

A typical estimation problem consists of the inverse problem of finding the vector parameter q , given observations $\{\xi_j\}$ of the state (or components of the state) corresponding to known inputs g or g_i . A typical control problem (for fixed parameter values q) might consist of minimizing a given payoff or cost functional subject to (1), (2), or (3), over some admissible class of control functions g or g_i .

§2. Motivating Examples

I. The LN_2 Wind Tunnel

The liquid nitrogen wind tunnel (National Transonic Facility) currently being constructed at NASA Langley Research Center is a cryogenic wind tunnel for which the cost of liquid nitrogen alone is estimated at \$6.5 $\times 10^6$ per year of operation. The tunnel represents the latest advances in technology in that essentially independent control of Mach number and Reynolds number (~ temperature) is an anticipated feature. Schematically, the tunnel can be represented as in Figure 1.

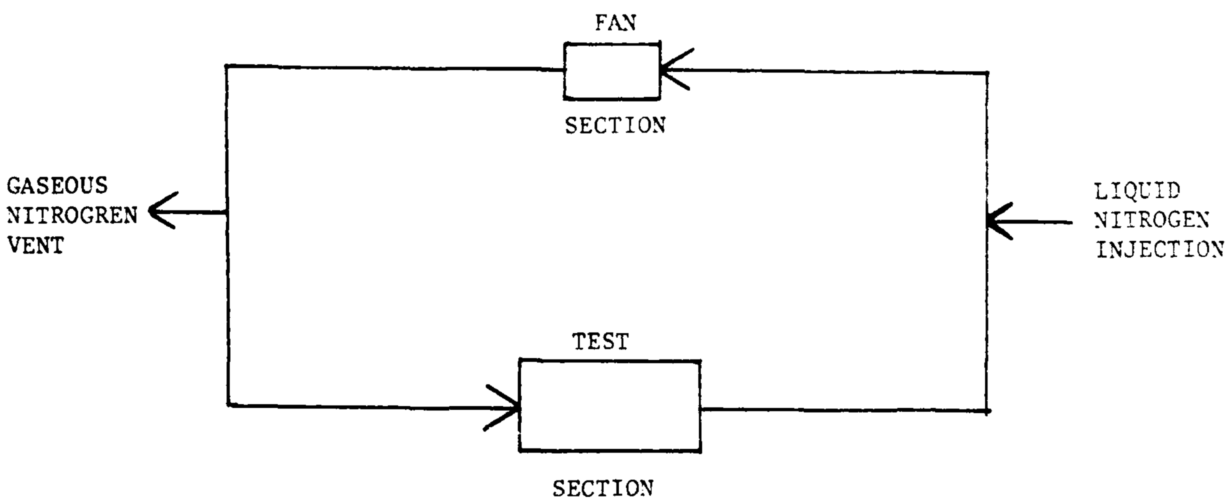

Figure 1

The basic physical model relating states such as Reynolds no., pressure, and Mach no. to controls such as LN_2 input, GN_2 bleed, and fan operation involves a formidable set of partial differential equations (Navier-Stokes) to describe fluid flow in the tunnel and test chamber. This model has, not surprisingly, proved to be very unwieldy computationally and probably cannot be used directly in design of sophisticated control laws. (Both open loop and feedback controllers are needed for efficient operation of the tunnel--and

this is clearly a desirable goal. Given the current estimates of costs of operation, the funds from only a 1 or 2% savings in operation costs would support a nontrivial amount of related research by scientists and engineers!!)

In view of the schematic in Figure 1, it is not surprising that engineers (e.g. see [18]) have proposed lumped parameter models (the variables representing values of states and controllers at various discrete locations in the tunnel) with transport delays to represent flow times in sections of the tunnel. A specific example is the model (see [1]) for the Mach no. (in the test chamber) loop which to first order is controlled by the fan guide vane angle setting (in the fan section)--i.e. $M(t) \sim GVA(t-\tau)$ where τ represents a transport time from the fan section to the test section.

In addition to the design of both open loop and closed loop controllers, parameter estimation techniques will be useful once data from the completed tunnel is available (current investigations involve use of data from a $\frac{1}{3}$ meter scale model of the tunnel).

II. Enzyme Tubular Reactors

Column reactors in which enzyme mediated chemical reactions take place to produce a desirable product (or products) from a given substrate (or substrates) are of some importance because of the numerous potential applications in commercial production (e.g., purification of fruit juices, proteolytic treatment of beer, synthesis of antibiotics and steroids). Research on the operation of such reactors has been carried out in the laboratory of D. Thomas at Universitè de Technologie de Compiègne for several years. Mathematical models for these processes (which involve reaction, diffusion, and convective transport) range from simple plug-flow (PF) models to full-fledged diffusion-convection-reaction (DCR) models [16], [19]. In an attempt to formulate models with the desirable accuracy exhibited by the DCR models (which are

computationally expensive and unwieldy to use, especially on small computers) but which approach the simplicity of the PF models (which in these applications prove too inaccurate in their representation of qualitative phenomena to be of practical use), J. P. Kernevez and his colleagues have proposed lumped parameter models with delays. In these models there are several delays representing convective transport and a number of diffusive transport mechanisms. One version of such models, which are nonlinear due to certain reaction velocity terms, is discussed in some detail by P. Daniel in [12] where additional references may also be found. To investigate the accuracy and potential usefulness of these models, efficient methods for parameter estimation (unknown parameters include several delays as well as kinetic constants) and control techniques for nonlinear delay systems are essential.

III. Gas and Oil Exploration and Recovery

a) Reservoir Engineering Problems

The importance of inverse or parameter estimation problems in the gas and oil industry is rather well-documented. One class of problems [8], [15], [26] involves use of the flow equations in a porous medium (a reservoir or oil/gas field) to determine the field porosity ϕ (the ratio of pore volume to total volume) and field permeability function k. A greatly simplified model would be based on an equation (derived from conservation of mass and Darcy's law--see [11], [20]) for the pressure $p = p(t,x,y)$ in a vertically homogeneous field of depth h, say

$$\phi ch \frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(\frac{hk}{\mu}\frac{\partial p}{\partial x}\right) + f ,$$

where μ = fluid viscosity, c = fluid compressibility, and f is a general sink/source term. The field usually contains a number of wells (for production

or observation or both) and a typical problem is to estimate ϕ and k or, alternatively, the total pore volume $\Phi \equiv \iint \phi h$, from observations of p at the well heads.

More realistic models involve several miscible fluids and a coupled set of partial differential equations [14], but the fundamental inverse problem is similar, only much more complicated, of course.

b) Seismic Exploration

A second class of inverse problems concerns determination of the elastic properties of an inhomogeneous medium via surface observations after perturbing "shocks" have produced waves in the medium. Models usually involve the equations of elasticity [2], [17]; for example, in the 1-dimensional problem one might consider

$$\rho(z)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial z}\left(E(z)\frac{\partial u}{\partial z}\right)$$

where ρ is the mass density and $E = \lambda + 2\mu$ for compressional or P-waves, $E = \mu$ for shear or S-waves with λ,μ the Lamé parameters. The boundary conditions at $z = 0$ (here z is the vertical distance from the surface) include excitation or perturbation of the medium (often this source input itself is a quantity to be "identified"). From observations at the surface $z = 0$ (these observations usually involve the unknown source input and a velocity term $\frac{\partial u}{\partial t}$ for particle displacement), one wishes to determine the unknown functions ρ and E and, in addition, the source term if it is unknown.

IV. Large Space Structures

Another class of control and identification problems for which the models are based on the equations for elastic structures are those dealing with

large space antennas. One such antenna that is currently being developed by NASA is the Maypole Hoop/Column antenna which is depicted in Figure 2.

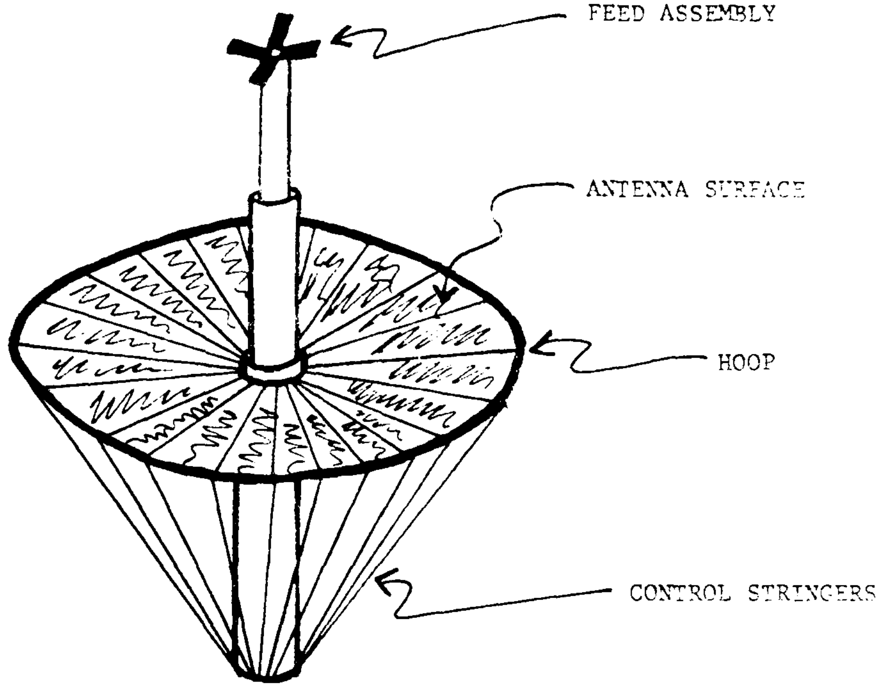

Figure 2

This antenna, which when fully deployed somewhat resembles an inverted umbrella (100 m. in diameter), consists of a membraneous surface of gold-plated molybdenum reflective mesh, a collapsible hoop or ring on which the surface is stretched, and a telescoping column to which the antenna surface is anchored and on which feed assemblies are mounted. The antenna in collapsed configuration (similar to the popular "travel umbrellas" that collapse to fit into a briefcase or small suitcase) is to be transported into space in the space shuttle; it is then deployed for use as a communication antenna. The antenna surface itself is flexible and its shape (and hence focusing properties) can be changed via control stringers attached to 48 equally spaced radial teflon coated graphite "cords" (4 control stringers per radial cord or "gore" edge). In addition to the dynamic identification and

control problems associated with initial deployment of the Maypole Hoop/Column, it is anticipated that after long periods of operation, the reflector surface will (due to changes in elastic properties and forces) require adjustment. Thus a static problem of interest consists of the following: Determine from observations (through sensing devices placed on the gore edges on the surface) the present configuration of the antenna surface and then effect the desired configuration or "displacement" through adjustment of the control stringers.

A typical problem then might involve a partial differential equation for the displacement of a circular membrane or thin plate, say $P(D,q)u = f$, where D represents spatial differential operators, q represents elastic parameters to be estimated, and f entails applied forces. A simple example might be

$$\frac{\partial}{\partial r}\left(r E \frac{\partial u}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{E}{r} \frac{\partial u}{\partial \theta}\right) = f$$

where $E = E(r,\theta)$.

V. Dispersion Models in Ecology

An important problem to population ecologists [21], [22] concerns the movement of insects (or, more generally, herbivores) through vegetation patches. Outbreaks or cyclical population explosions of some insects are observed and it is believed that the nature of the transport mechanisms for the insects affect the occurrence (or lack thereof) of outbreaks and their magnitude and periodicity. Typical equations to describe movement of the insects involve both diffusive and advective (convective) terms, e.g. for 1-dimensional models

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x}(VN) = \frac{\partial}{\partial x}\left(D \frac{\partial N}{\partial x}\right) + f(N) ,$$

in addition to the usual sink/source terms f . Depending on the species involved, it is generally expected that D and/or V can depend on N , the population level, and/or x , the spatial variable. The diffusion coefficient D and the convective velocity V may alternatively, or, in addition, depend on temperature or time (e.g. as in seasonal migration of pests).

There are numerous estimation and control problems of importance in the context of ecological investigations. Typically one wishes to determine the coefficient functions D and V from observations of N and once this is done, one might wish to estimate the optimal vegetation density in a patch in order to hold population levels in the patch to a minimum, or at least below some given level.

VI. Transport Models in Physiology

In physiology a great deal of research is devoted to questions concerning transport mechanisms such as simple (passive) diffusion, bulk flow or convective transport, facilitated diffusion, and active transport. An example is the effort [9], [10], [23], [24] devoted in recent years to the controversy involving bulk flow vs. molecular diffusion of brain interstitial fluid in gray and white matter. The mathematical models again are based on the convection-diffusion equation

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} = D \frac{\partial^2 u}{\partial x^2}$$

for the concentration u of a labeled substance such as sucrose in brain tissue. From experimental data (for u at various times and locations in the tissue samples) one seeks to estimate values for V and D in gray and white matter and contrast the transport properties of each type of tissue. We shall, in a subsequent section of this presentation, discuss in some detail an

application to these transport problems of some of the methods that are the focus of attention in this lecture.

§3. Theoretical Foundations

We turn now to a discussion of the theoretical techniques that one can employ to establish convergence results for certain approximation schemes for nonlinear systems such as (1), (2) or (3). For the sake of brevity we shall restrict our considerations to parameter estimation problems. A discussion of the use of the ideas presented here in control problems can be found in [12] in the case of nonlinear delay systems while the case of distributed parameter control problems is considered briefly in [6].

For the purposes of illustration we shall use a least squares formulation (for a discussion of maximum likelihood estimator ideas, see [3] and the references therein) of the parameter estimation problem. In particular, one seeks to minimize

$$J(q) = \frac{1}{2}\sum_{i=1}^{m} |y(t_i;q) - \xi_i|^2$$

over a given set Q of admissible parameters. Here ξ_i is an observation for the output $t \to y(t;q)$ at t_i with $y(t) = \mathcal{C}x(t)$ in the case of (1), $y(t) = \mathrm{col}(\mathcal{C}u(t,x_1),\ldots,\mathcal{C}u(t,x_p),\mathcal{D}u_t(t,x_1),\ldots,\mathcal{D}u_t(t,x_p))$ in the case of (2), and $y(t) = \mathrm{col}(\mathcal{C}u(t,x_1),\ldots,\mathcal{C}u(t,x_p))$ in the case of (3), where $\mathcal{C}$ and $\mathcal{D}$ are matrix operators of appropriate dimension and rank.

Our approach entails rewriting (1), (2), or (3) as an abstract equation

$$\begin{aligned} \dot{z}(t) &= \mathcal{A}(q)z(t) + G(t) \\ z(0) &= z_0 \end{aligned} \tag{4}$$

in an appropriately chosen Hilbert space Z. The operator $\mathcal{A}$ may be linear or nonlinear and depends on the unknown parameters q. We reformulate the estimation problems as ones of minimizing

$$J(q) = \frac{1}{2} \sum_{i=1}^{m} |\Gamma(z(t_i;q)) - \xi_i|^2$$

where $y(t;q) = \Gamma(z(t;q))$ is an appropriately defined output.

We take a classical Ritz-Galerkin type approach to reducing these infinite-dimensional state space problems to a sequence of approximating finite dimensional state space problems that are readily solved numerically. For a given sequence Z^N of "subspaces" of Z with "projections" $P^N: Z \to Z^N$, we minimize

$$J^N(q) = \frac{1}{2} \sum_{i=1}^{m} |\Gamma(z^N(t_i;q)) - \xi_i|^2$$

over Q where z^N is the solution of an approximating system

$$\begin{aligned} \dot{z}^N(t) &= A^N(q)z^N(t) + P^N G(t) \\ z^N(0) &= P^N z_0 \,. \end{aligned} \tag{5}$$

In the methods discussed here we always take $A^N(q) = P^N A(q) P^N$ and obtain a state convergence $z^N(t;q) \to z(t;q)$. The ultimate goal, of course, is to insure convergence of some sequence $\{\bar{q}^N\}$ of solutions of approximating estimation problems involving (5) to a solution $\bar{q}$ of the problems involving (4). This objective can be attained in the cases of the "modal" and "spline" schemes we have developed and tested numerically in [4], [5], [6], [7], [12].

To date we have employed two different theories to establish state and parameter convergence. For distributed parameter systems, both modal [6] and spline [7] schemes have been investigated using an abstract semigroup formulation and Trotter-Kato type theorems. Briefly, one establishes that the linear operators A and A^N (we suppress the q dependence here) satisfy a

uniform (in q and N) dissipativeness condition and generate C_0-semigroups $T(t)$ and $T^N(t)$ respectively. Then treating the nonlinearities $G(\sigma) = F(q,\sigma,z(\sigma))$ (F is defined in an appropriate manner using f from (2) or (3)) as perturbations, one considers in place of (4) and (5) the implicit equations

(6) $$z(t) = T(t)z_0 + \int_0^t T(t-\sigma)F(q,\sigma,z(\sigma))d\sigma$$

and

(7) $$z^N(t) = T^N(t)P^N z_0 + \int_0^t T^N(t-\sigma)P^N F(q,\sigma,z^N(\sigma))d\sigma \ .$$

The basic tool then is a Trotter-Kato type result which, under the conditions:

(8i) $|T^N(t)| \leq Me^{\omega t}$ for some M and ω independent of N ;

(8ii) there exists a set $\mathcal{D} \subset \mathrm{Dom}(A)$, $\mathcal{D}$ dense in Z , such that $(\lambda_0 - A)\mathcal{D}$ is dense in Z for some $\lambda_0 > 0$;

(8iii) $|A^N z - Az| \to 0$ for $z \in \mathcal{D}$;

guarantees the convergence

(8iv) $T^N(t)z \to T(t)z$ for $z \in Z$, uniformly in t .

The convergence in (8iv), along with (6), (7) and $P^N \to I$ strongly, can be used to argue state convergence $z^N(t) \to z(t)$. This in turn can be used to establish a desired parameter convergence (for the rather technical details-- which are nontrivial when the full dependence of the operators, projections, semigroups and, in some cases, the subspaces, on the unknown parameters q is

taken into account--one should consult [6] and [7]).

A somewhat different approach to spline methods for delay systems (') has been taken in [4], [5], [12] where the nonlinearity f is treated directly as part of a nonlinear operator $A = A(t)$ (which is now possibly time dependent). In this case one uses the implicit equations

$$z(t) = z_0 + \int_0^t \{A(\sigma)z(\sigma) + G(\sigma)\}d\sigma$$

$$z^N(t) = P^N z_0 + \int_0^t \{A^N(\sigma)z^N(\sigma) + P^N G(\sigma)\}d\sigma$$

in place of (4) and (5). Under reasonable conditions on f one can establish dissipative type inequalities

$$\langle\langle A^N(\sigma)z - A^N(\sigma)w, z - w\rangle\rangle \leqslant \omega(\sigma)\langle\langle z - w, z - w\rangle\rangle$$

where $\langle\langle\ ,\ \rangle\rangle$ is a specially defined inner product on Z. With some elementary analysis and use of a Gronwall inequality, one then obtains estimates for $|z^N(t) - z(t)|$ in terms of integrals of $\{A^N(\sigma) - A(\sigma)\}z(\sigma)$. Desired convergence results then follow from convergence properties of A^N. Again the technical details become quite involved when one treats general non-linear delay systems with multiple unknown delays. These can be found in [5], [12].

We remark that one need not have Z^N a subspace of Z in carrying out the above theories. Indeed in both cases (delay systems with unknown delays, distributed parameter systems with unknown coefficients) outlined above, one finds that the appropriate P^N, Z^N, and Z all depend on the unknown parameters q (through the domain of the function space in the case of unknown delays and through the inner product for Z and Z^N in the case of some distributed parameter examples as well as the unknown delays problems) which of course

vary as one iterates through the sequence of approximating problems (i.e., on N). This feature results in interesting difficulties from both a conceptual and computational viewpoint.

§4. An Application to Transport in Brain Tissue

We return to the example VI of §2 involving the transport of labeled sucrose in gray and white matter. A complete description of the experimental procedures and the questions being investigated can be found in [24]. Briefly, cats are anesthetized and experiments of either 1, 2, or 4 hours duration are carried out. Labeled sucrose is perfused into the lateral ventricle. At the end of the perfusion period, the animals are sacrificed and their brains are removed and frozen. Well-stained areas of gray and white matter perpendicular to the ventricular surface (along the x-axis in our notation below) are sectioned and analyzed. This yields data corresponding to a fixed time t_i for a maximum of 4 spatial locations $x_j, j = 1,\ldots,4$, in gray matter and 8 spatial locations x_j in white matter. From this data $\{\hat{u}(t_i,x_j)\}$ for the concentration u, one wishes to compare transport in gray matter with that in white matter. The primary questions pertain to transport via molecular diffusion alone vs. transport via diffusion and bulk (convective) flow. In particular, the mathematical problems reduce to those of estimating D, V, and C_0 in

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} = D \frac{\partial^2 u}{\partial x^2}$$

$$u(t,0) = C_0 \ .$$

In the early experimental work, data for only one time t_1 (1, 2, or 4 hours) and for anywhere from 4 to 8 spatial locations x_j were available. A substantial concern is whether one can develop accurate methods for estimation of the three parameters in question from such limited data.

We have successfully applied the modal methods of Example 4.4 of [6] to investigate these problems. We first summarize briefly the pertinent ideas behind the methods. For the purpose of illustration, consider the example

$$u_t = q_1 u_{xx} + q_2 u_x$$

$$u(t,0) = u(t,1) = 0$$

$$u(0,x) = \phi(x) \ ,$$

which can be reformulated in the form (4) in $Z = L_2(0,1)$ by choosing $A(q) = q_1 D^2 + q_2 D$ (here D is the differential operator in $L_2(0,1)$) with $\mathrm{Dom}(A(q)) = H^2 \cap H_0^1$. It can be argued that $A(q)$ is uniformly maximal dissipative (although it is not self-adjoint). For the development of modal approximation schemes, general spectral results given in [13] can be employed. It can be seen that $A(q)$ is a relatively bounded perturbation of a discrete spectral operator and is itself a discrete spectral operator with a resolution of identity and associated eigenmanifolds and projections. The eigenvalues are found to be $\lambda_j(q) = -j^2\pi^2 q_1 - q_2^2/2q_1$ with associated eigenfunctions $\psi_j(q) = \exp\{-q_2 x/2q_1\}\sin j\pi x$. The natural modes or eigenfunctions form a complete (but not orthogonal) set in $Z = L_2(0,1)$. However a choice of $\tilde{Z}^N = \mathrm{span}\{\psi_1(q),\ldots,\psi_N(q)\}$, while desirable theoretically, is not useful in parameter estimation algorithms since the basis elements are then dependent upon the unknown parameters (and thus change with each new estimate of the q's). One can use instead the near-modal functions $\phi_j(x) = \sqrt{2}\sin j\pi x$ and take $Z^N = \mathrm{span}\{\phi_1,\ldots,\phi_N\}$ with, of course, $A^N = P^N A P^N$ where P^N is the canonical projection of Z onto Z^N .

Convergence can be argued using the Trotter-Kato formulation of (8i) - (8iv) above. The stability condition (8i) follows immediately from the uniform dissipativeness. Choosing $\mathcal{D} = \bigcup_{N=1}^{\infty} \tilde{Z}^N(\bar{q})$, where $\bar{q}$ is a limit of the sequence of estimates $\bar{q}^N$, the spectral results yield (8ii) trivially while one must work somewhat more to establish (8iii).

With regard to implementation, the scheme offers some nice computational features since the matrix realizations of the operators $A^N(q)$ are given by

$$[A^N(q)]_{ij} = \begin{cases} -q_1 i^2\pi^2 & i = j \ . \\ 0 & i \neq j \ , \ i + j \ \text{ even,} \\ 2jq_2[\frac{2i}{i^2-j^2}] & i \neq j \ , \ i + j \ \text{ odd.} \end{cases}$$

Turning to our investigation of these methods for possible use in the brain transport questions, we first tested the methods with an example for which the solution was "known". (J. Crowley and M. Garrett carried out the computations for this problem. J. Saltzman supplied a "known" solution technique involving an infinite series which was used to generate "data" corresponding to fixed parameter values in the equation. This technique is totally unrelated to the methods we were testing.) The example used was

$$u_t - q_2 u_x = q_1 u_{xx}$$

$$u(t,0) = q_3 \ , \ u(t,1) = 0$$

$$u(0,x) = \phi(x) \ ,$$

where $\phi(x) = a_0 x^2 + a_1 x + a_2$ is a quadratic satisfying $\phi(0) = 1$, $\phi(1) = 0$, and $\max \phi = \phi(\frac{1}{4})$. "Data" were generated corresponding to true values $q_1^* = .3$, $q_2^* = 1.75$, and $q_3^* = 1.0$. A number of numerical trials with the above described "modal" scheme were conducted in which the inverse problem for varying amounts of "data" was "solved." We summarize briefly some of our findings. In the examples presented here, the notation $I = k$, $J = p$ in an example indicates that the data set for this test consisted of values $u(t_i,x_j)$, $i = 1,\ldots,k$, $j = 1,\ldots,p$ (i.e., $k \times p$ "observations" were employed in the inverse problem).

<u>Example 1</u>: It was assumed that q_3 was known and an attempt to fit the data was made by searching for q_1 and q_2 . Initial guesses (for each value of N

tried) were $q_1^{N,0} = 1.0$, $q_2^{N,0} = 0.0$. The "converged" values (corresponding to $N = 8$ or 16 in this and all the examples presented here) were:

$$I = 1,\ J = 3:\quad \bar{q}_1 = 1.82 \quad \bar{q}_2 = .83$$

$$I = 2,\ J = 3:\quad \bar{q}_1 = .2979 \quad \bar{q}_2 = 1.7557 \ .$$

Example 2: This example was exactly the same as Example 1 except initial guesses $q_1^{N,0} = .75$, $q_2^{N,0} = 1.0$ (somewhat closer to the "true" values than those used in Ex. 1) were used. For $I = 1,\ J = 3$, the results were $\bar{q}_1 = .3009$, $\bar{q}_2 = 1.7529$, quite acceptable in this case.

Example 3: We investigated the effect of using increasingly more spatial points in our data grid (*i.e.*, $I = 1$ *with* $J = 4,5,6$). *For initial guesses* $q_1^{N,0} = .8$, $q_2^{N,0} = .9$, increasingly better estimates were obtained as the number of spatial grid points increased. We obtained:

$$I = 1,\ J = 3:\quad \bar{q}_1 = .6115 \quad \bar{q}_2 = 1.4903$$

$$I = 1,\ J = 4:\quad \bar{q}_1 = .3018 \quad \bar{q}_2 = 1.7468$$

$$I = 1,\ J = 5:\quad \bar{q}_1 = .2978 \quad \bar{q}_2 = 1.7492$$

$$I = 1,\ J = 6:\quad \bar{q}_1 = .2984 \quad \bar{q}_2 = 1.7493 \ .$$

It is clear that 4 points in the spatial grid yields an adequate amount of data for this example.

Example 4: In this case we wished to estimate q_1, q_2 and the boundary concentration q_3 . Initial guesses were $q_1^{N,0} = .8$, $q_2^{N,0} = .9$, $q_3^{N,0} = .5$. The converged values were:

$$I = 1,\ J = 6:\quad \bar{q}_1 = .2990 \quad \bar{q}_2 = 1.6983 \quad \bar{q}_3 = 1.0356$$

$$I = 2,\ J = 6:\quad \bar{q}_1 = .2997 \quad \bar{q}_2 = 1.7469 \quad \bar{q}_3 = 1.0012\ .$$

Example 5: As a final test we modified the initial function ϕ used in the above examples to $\phi(0) = 1$, $\phi(x) = 0$ for $x \neq 0$. This represents the type problem (one with a discontinuity in the boundary-initial data) that one encounters when using the actual data collected in the experiments with cats described above. Again the results obtained were encouraging. With $I = 2$, $J = 6$ and initial guesses $q_1^{N,0} = .8$, $q_2^{N,0} = .9$, converged values of $\bar{q}_1 = .3019$, $\bar{q}_2 = 1.7635$ were found with a residual sum of squares of 4.3×10^{-3} .

In summary, the numerical tests reveal that it is probably unreasonable to expect to solve with the "modal" methods the inverse problems for $I = 1$, $J = 3$ in most cases. However, problems with $I = 2$, $J = 3$ correspond to a reasonable number of spatial observations for the method in some cases. (By changing labels during the perfusion period, Rosenberg, Kyner, and colleagues are now collecting data with two time grid points.) For data from white matter (where $J = 6$ is feasible), the methods should prove useful in estimating D , V , and C_0 in the transport models.

We have, in fact, used the methods with actual data sets $(I = 1, J = 8)$ for white matter supplied by Kyner and his associates. The "modal" methods appear to consistently perform in an acceptable manner. Typical values obtained in solving the inverse problems are $D = 2.7 \times 10^{-6}\ \mathrm{cm}^2/\mathrm{sec.}$, $V = -5.99\ \mu\mathrm{m/min.}$, $C_0 = 128.5$, values that are consistent with expectations based on values obtained by Kyner and associates using other techniques.

We anticipate that extensions of these methods (or perhaps the spline

methods developed in [7]) will prove useful in future investigations of the channeling structure in white matter (in these problems the velocity coefficient V will be spatially dependent as will also, in some cases, the coefficient of diffusion D).

§5. Estimation of Function Space Parameters

The theory developed in [6] and [7] deals with estimation of parameters in Euclidean space sets. The framework is, however, general enough to allow one to treat problems in which unknown function space parameters must be estimated. In this section of our presentation we shall give a brief sketch of how one further develops such a theory. At the same time we shall illustrate some of the ideas fundamental to spline methods as opposed to the "modal" methods discussed earlier.

In order to demonstrate the ideas we shall consider an equation of the porous media type (see §2.IIIa)); that is, we consider

$$q_1(x)\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(q_2(x)\frac{\partial u}{\partial x}\right) + f \tag{9}$$

with homogeneous boundary conditions $u(t,0) = u(t,1) = 0$. In relating this to the porous media application (then u = pressure), one might consider large fields for which the boundary terms are either constant or slowly varying in time. In either case, such nonhomogeneous boundary problems can be transformed to problems with homogeneous boundary conditions in a quite standard manner. With certain smoothness assumptions on q_1, q_2, the operator in equation (9) can be viewed as a standard Sturm-Liouville operator (i.e. identify $q_1 \sim k$, $q_2 \sim p$ in the usual notation for the coefficient functions--see (3) above and p. 40-42 of [6]). For our discussions here we shall assume that $q = (q_1,q_2)$ is to be chosen from a parameter space $Q \subset L_2(0,1) \times L_2(0,1)$ satisfying $Q \subset \{(q_1,q_2) \in H^2 \times H^3 | q_2 > 0 , 0 < m \leq q_1 \leq M\}$. (The smoothness hypothesized will guarantee certain smoothness properties for the eigenfunctions to be discussed momentarily.)

We rewrite (9) as an equation

$$\dot{z}(t) = A(q)z(t) + F(q,f) \tag{10}$$

in the state space $Z = X(q) = L_2(0,1)$ where we take as inner product $\langle\phi,\psi\rangle_q = \int_0^1 \phi\psi q_1$. (Here the spaces do depend on the unknown parameters, a complicating possibility we mentioned earlier.) The operators in (10) are given by $F = \frac{1}{q_1} f$ and $A(q)\psi = \frac{1}{q_1} D(q_2 D\psi)$, where $\mathrm{Dom}(A(q)) = H^2 \cap H_0^1$.

Simple integration by parts arguments yield $\langle A(q)z,z\rangle_q \leq 0$ so that $A(q)$ is uniformly dissipative in $X(q)$. In fact $A(q)$ is maximal dissipative and generates a C_0-semigroup, and we are thus in a position to consider (6), (7) and the Trotter-Kato approach to approximation schemes.

To describe the spline methods we need to recall the definition of some standard cubic spline basis elements. For any positive integer N we let $t_j^N = j/N$, $j = -3,\ldots,N+3$, and let $\tilde{B}_j^N$, $j = -1,\ldots,N+1$, be the cubic spline that vanishes outside (t_{j-2}^N, t_{j+2}^N) , has value 4 and slope 0 at t_j^N , value 1 and slope $3N$ at t_{j-1}^N , and value 1 and slope $-3N$ at t_{j+1}^N . (See [25], p. 73--and note that our elements here differ from those of Schultz only by a multiplicative factor of 24 .)

For our modified basis elements B_j^N we take the restriction to $[0,1]$ of the following:

$$
\begin{aligned}
B_0^N &= \tilde{B}_0^N - 4\tilde{B}_{-1}^N \\
B_1^N &= \tilde{B}_0^N - 4\tilde{B}_1^N \\
B_j^N &= \tilde{B}_j^N , \quad j = 2,\ldots,N-2 , \\
B_{N-1}^N &= \tilde{B}_N^N - 4\tilde{B}_{N-1}^N \\
B_N^N &= \tilde{B}_N^N - 4\tilde{B}_{N+1}^N .
\end{aligned}
$$

We note that these elements are in $\mathrm{Dom}(A(q))$.

We define our approximation subspaces $X^N(q) \subset X(q)$ by $X^N(q) = \text{span}\{B_0^N,\ldots,B_N^N\}$ and let $P^N(q)$ be the canonical projection of $X(q)$ onto $X^N(q)$, i.e., $P^N(q)\psi = \sum_{j=0}^{N} \langle\psi, B_j^N\rangle_q B_j^N$. Finally as usual we take $A^N = A^N(q) = P^N(q)A(q)P^N(q)$.

Under an assumption that Q is compact in $H^0 \times H^0$, one can argue in this case that solutions to the estimation problems for (5) (or (7)) do exist. We in fact assume that Q is compact in the $C \times H^1$ topology so that we henceforth assume without loss of generality (possibly by taking a subsequence) that we have a sequence $\{q^N\}$ of solutions to the estimation problems satisfying $q^N \to \bar{q}$ in $C \times H^1$ with $\bar{q} \in Q$.

We briefly indicate the steps to verify (8i) - (8iii) to insure convergence of the semigroups generated by $A^N(q^N)$ to the semigroup generated by $A(\bar{q})$. (As we have noted before, this is the fundamental convergence result needed for both state and parameter convergence.) The stability requirement (8i) follows from

$$\langle A^N(q^N)z,z\rangle_{q^N} = \langle A(q^N)P^N(q^N)z,\ P^N(q^N)z\rangle_{q^N} \leq 0\ ,$$

the inequality being a result of the uniform dissipativeness of $A(q^N)$.

The operator $A(\bar{q})$ has, by the usual spectral results, a CONS of eigenfunctions $\{\Psi_j(\bar{q})\}$. In (8ii) we take

$$\mathcal{D} = \bigcup_{N=1}^{\infty} \text{span}\{\Psi_1(\bar{q}),\ldots,\Psi_N(\bar{q})\}\ .$$

It is then easily seen that the conditions of (8ii) obtain from the completeness of the Ψ_j and the relationship $(\lambda_0 - A(\bar{q}))\Psi_j(\bar{q}) = (\lambda_0 - \lambda_j)\Psi_j(\bar{q})$.

We finally consider (8iii) and note that the spaces $X(q)$, $q \in Q$, are

all equivalent (recall $0 < \mu \leq q_1 \leq M$), a fact which plays a fundamental role in the basic theory developed in [6]. Indeed we may, in considering any convergence results, equivalently use the L_2 topology. Thus, to establish (8iii), it suffices to argue

$$A^N(q^N)\Psi_j \to A(\bar{q})\Psi_j \tag{11}$$

in L_2 . From the smoothness assumptions on Q (and hence $\bar{q}$) it is easily seen that $\Psi_j \in H^4$. Since $\Psi_j \in \mathrm{Dom}(A(\bar{q}))$ we also have $\Psi_j \in H_0^1$. Hence, it suffices to fix $\psi = \Psi_j(\bar{q})$ in $H^4 \cap H_0^1$ and argue that (11) holds whenever $q^N \to \bar{q}$ in $C \times H^1$.

Estimates similar to those we need can be found in Theorem 6.13, p. 82 of [25]. However we cannot use those estimates directly since our projections $P^N(q^N)$ (onto $X^N(q^N)$) are not the same as the standard projections (of Thm. 6.13) of L_2 onto $S(N) = \mathrm{span}\{\tilde{B}_{-1}^N, \tilde{B}_0^N, \ldots, \tilde{B}_{N+1}^N\}$. But using fundamental ideas similar to those found in [25] (e.g., the Schmidt inequality and estimates for the appropriate interpolating splines) one can establish:

For $\psi \in H^4 \cap H_0^1$,

$$|\psi - P^N\psi| \leq \frac{K_1}{N^4} |D^4\psi|$$

$$|D(\psi - P^N\psi)| \leq \frac{K_2}{N^3} |D^4\psi|$$

$$|D^2(\psi - P^N\psi)| \leq \frac{K_3}{N^2} |D^4\psi|$$

where the norms are the usual L_2 norm , $P^N = P^N(q^N)$ as defined earlier, and the constants K_1, K_2, K_3 are independent of N and ψ .

We thus have $D^2\psi^N \to D^2\psi$ in L_2 , $D\psi^N \to D\psi$ in C , where $\psi^N = P^N\psi$. Furthermore we know that $q_1^N \to \bar{q}_1$ in C , $q_2^N \to \bar{q}_2$ in H^1 . Since $P^N \to I$ it thus follows from elementary arguments that

$$A^N(q^N)\psi = P^N\{(1/q_1^N)Dq_2^N D\psi^N + (q_2^N/q_1^N)D^2\psi^N\}$$

converges in L_2 to

$$A(\bar{q})\psi = (1/\bar{q}_1)D\bar{q}_2 D\psi + (\bar{q}_2/\bar{q}_1)D^2\psi \ .$$

The assumptions on Q made above are not unreasonable for many applications. We have successfully used these spline methods in computational packages for function space parameter estimation in models for insect dispersion (see §2.V). In those applications, the smoothness and compactness assumptions listed above are satisfied when one formulates the problems and parameterizes Q in a way that is natural for and consistent with the experimental and theoretical efforts of population ecologists.

REFERENCES

[1] E. A. Armstrong and J. S. Tripp, An application of multivariable design techniques to the control of the National Transonic Facility, NASA Tech. Paper 1887, NASA-LRC, August, 1981.

[2] A. Bamberger, G. Chavent and P. Lailly, About the stability of the inverse problem in 1-D wave equations--application to the interpretation of seismic profiles, Appl. Math. Optim. 5 (1979), 1-47.

[3] H. T. Banks, Parameter identification techniques for physiological control systems, in Mathematical Aspects of Physiology (F. Hoppensteadt, ed.) Vol. 19, Lec. in Applied Math, Amer. Math. Soc., 1981, pp. 361-383.

[4] H. T. Banks, Indentification of nonlinear delay systems using spline methods, in Proc. Intl. Conf. on Nonlinear Phenomena in Math. Sciences (V. Lakshmikantham, ed.), Academic Press, 1981, to appear.

[5] H. T. Banks and P. L. Daniel, Estimation of delays and other parameters in nonlinear functional differential equations, to appear.

[6] H. T. Banks and K. Kunisch, An approximation theory for nonlinear partial differential equations with applications to identification and control, LCDS Tech. Rep. 81-7, Brown University, April, 1981.

[7] H. T. Banks, J. M. Crowley, and K. Kunisch, Cubic spline approximation techniques for parameter estimation in distributed systems, to appear.

[8] G. Chavent, M. Dupuy and P. Lemonnier, History matching by use of optimal control theory, Soc. Petroleum Eng. J., (1975), 76-86.

[9] H. F. Cserr, D. N. Cooper and T. H. Milhorat, Flow of cerebral interstitial fluid as indicated by the removal of extracellular markers from rat caudate nucleus, Exp. Eye Res. Suppl., 25 (1977), 461-473.

[10] H. F. Cserr, D. N. Cooper, P. K. Suri, and C. S. Patlak, Efflux of radiolabeled polyethylene glycols and albumin from rat brain, Am. J. Physiol. 240 (Renal Fluid Electrolyte Physiol. 9), 1981, F319-F328.

[11] L. P. Dake, Fundamentals of Reservoir Engineering, Elsevier Scientific Publ. Co., 1978.

[12] P. L. Daniel, Spline-based approximation methods for the identification and control of nonlinear functional differential equations, Ph.D. Thesis, Brown University, 1981.

[13] N. Dunford and J. T. Schwartz, Linear Operators: Part III, Spectral Operators, Wiley-Interscience, New York, 1971.

[14] R. E. Ewing and M. F. Wheeler, Galerkin methods for miscible displacement problems in porous media, SIAM J. Num. Anal. 17 (1980), 351-365.

[15] G. R. Gavalas and J. H. Seinfeld, Reservoirs with spatially varying properties: estimation of volume from late transient pressure data, Soc. Petroleum Eng. J. (1973), 335-342.

[16] G. Gellf and J. Henry, Experimental and theoretical study of diffusion, convection and reaction phenomena for immobilized enzyme systems, in Analysis and Control of Immobilized Enzyme Systems, (D. Thomas and J. P. Kernevez, eds.) North-Holland/American Elsevier, New York, 1976, pp. 253-274.

[17] F. Grant and G. West, Interpretation Theory in Applied Geophysics, McGraw-Hill, 1965.

[18] G. Gumas, The dynamic modeling of a slotted test section, NASA CR-159069, 1979, Penn State University.

[19] J. Henry, Contrôle d'un réacteur enzymatique à l'aide de modèles à paramètres distribués: quelques problèmes de contrôlabilité de systemes paraboliques, Thèses d'État, Université Paris VI, 1978.

[20] M. K. Hubbert, Darcy's law and the field equation of the flow of underground fluids, Petroleum Transactions, AIME, 207 (1956), 222-239.

[21] Peter Kareiva, The application of diffusion to herbivore models, Ecological Monograph, to appear.

[22] A. Okubo, Diffusion and Ecological Problems: Mathematical Models, Biomathematics, Vol. 10, Springer, New York, 1980.

[23] G. A. Rosenberg and W. T. Kyner, Gray and white matter brain-blood transfer constants by steady-state tissue clearance in cat, Brain Res. 193 (1980), 59-66.

[24] G. A. Rosenberg, W. T. Kyner, and E. Estrada, Bulk flow of brain interstitial fluid under normal and hyperosmolar conditions, Am. J. Physiol. 238 (Renal Fluid Electrolyte Physiol. 7), 1980, F42-F49.

[25] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, 1973.

[26] P. C. Shah, G. R. Gavalas and J. H. Seinfeld, Error analysis in history matching: the optimum level of parametrization, Soc. Petroleum Eng. J., (1978), 219-228.